Resolution of Resonances in a General Purpose Quantum Device Simulator (NEMO)

GERHARD KLIMECK[a,b,*], ROGER K. LAKE[b], R. CHRIS BOWEN[a],
CHENJING L. FERNANDO[a,b] and WILLIAM R. FRENSLEY[a]

[a]*School of Engineering, University of Texas at Dallas, Richardson, TX 75083-0688,* [b]*Corporate R&D, Texas Instruments Incorporated, Dallas, TX 75235*

Electron transport in quantum devices is governed by discrete quantum states due to electron confinement. A crucial requirement for the modeling of quantum devices is the the numerical identification and resolution of these quantum states. We present an algorithm utilized in our general purpose quantum device simulator (NEMO), where we locate the resonances of the system first and then generate the optimized grid used to integrate over the resonances. We find this algorithm important in the modeling of coherent transport involving ultrafine resonances and crucial for the modeling of incoherent transport.

Keywords: Resonant tunneling, quantum device, Numerical Integration, Grid generation, NEMO

One major problem in the development of a high bias current-voltage simulator for quantum devices is the resolution and treatment of resonant states. The resonance states strongly influence the transmission through and the scattering in the device. To calculate observables like the current and the charge at high bias we have to integrate in energy (and/or momentum) over these resonances. These resonance states are as sharp as 1neV-5meV in an energy range of 0-1000meV spanned by the applied bias. A homogeneous energy grid resolution of 1μeV over an energy range of 1000meV would require 1,000,000 nodes which is unfeasible for our purposes, where a single integration point may be very expensive to compute, or the storage of all integration points is necessary. Since there are only few sharply varying features the numerical analysis should be carried out on a well optimized inhomogeneous energy grid. In addition to the sharp quantum features we have to resolve more

smoothly varying features such as the lead carrier distributions of 5meV-50meV with a temperature dependent modulation of 0.3meV-25meV. Figure (1a) illustrates these spectral features for the case of a typical resonant tunneling diode.

The standard [1] solution to this problem of integration over sharply varying integrands is the sucessive partitioning of the integration range. An adaptive grid is constructed by adding nodes within the integration range of interest as a test of the integration convergence. While this trial and error method is foolproof as long as the convergence criteria are stringent enough, it does have 2 conceptual drawbacks.

1. The optimization is based on a relative error analysis. I.e. the improvement of the integrand due to adding a node is measured in the small integration segment, rather than the overall integral. The requested relative accuracy is therefore not in direct relation to the absolute accuracy of the integral. The

---

* Present address. Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109-8099, USA
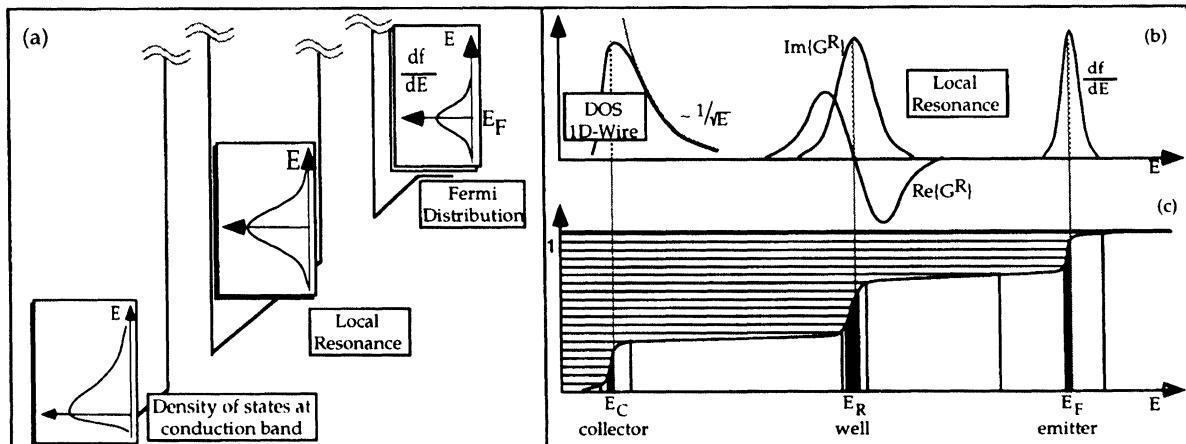
FIGURE 1 Construction of an inhomogeneous energy grid using the location of spectral features. (a) Schematic conduction band profile of a resonant tunneling diode. The inserted panels indicate the spatial and energetic location of spectral features that we wish to resolve: localized resonances, Fermi factors and conduction band edges. (b) Spectral features on one energy axis. In addition to (a) we show the real part of the retarded Green function (principal value integral of a Lorentzian). (c) The single spectral features in (b) can be integrated analytically to obtain a monotonic function. An optimized inhomogeneous energy grid can be obtained by projecting a homogeneous grid from the y-axis of the monotonic function onto the energy axis

procedure will resolve curvature features well, however it will also place many nodes in energy ranges where the total integral contribution is insignificant. An optimization procedure minimizing the absolute error is more desirable to avoid the calculation of insignificant integral contributions.

2. Nodes are added as a pure test whether they are needed for the integration or not. Adding nodes in an integration can complicate algorithms dramatically, if other quantities which depend on the integral and modify the integrand are to be updated as well. For vectorization purposes we prefer to deal with fixed grids.

The expensive task in the sucessive partitioning approach is the locating of the resonances. If we would know where the resonances were, we could conceive of a method to place an inhomogeneous grid smartly. Resonance finding algorithms for single band [2] and multi band [3] Hamiltonians have been developed and we use them as part of our fixed grid generation. These algorithms return the eigenvalues $(E_r - i\Gamma)$ of the single electron tight binding Hamiltonians describing the quantum device. The location, E_r, and width, Γ, of the resonant states in the device lays the foundation for the generation of a grid that resolves these features well.

For numerical stability in terms of round off errors we would like the integration contribution from one energy node to the next to be about the same. This suggests a very simple scheme which lets us optimize the energy grid with respect to integral quantities. The integral over a spiked function is a monotonic increasing function. Figure (1c) indicates how a homogeneous grid on the abscissa of a monotonic function projects to an inhomogeneous grid on the ordinate. The question now remains how we get the monotonic functions of the right shape.

Close to the resonance we can approximate the density of states $-\frac{1}{\pi}IM\{G^R\}$ where G^R is the retarded Green function [3] by a Lorentzian at energy E_r and width Γ. The real part of G^R is proportional to the principal value integral of that Lorentzian. The integral over the Lorentzian function and the magnitude of the principal value integral can be performed analytically resulting in monotonic functions $m_{IM}(E)$ and $m_{RE}(E)$. The real part of G^R is doubly peaked at $E_r \pm \Gamma/2$. We enhance the resolution of the resonance by placing two monotonic contributions m_{RE-IM} corresponding to $m_{IM}(E)$ at the energies $E_r \pm \Gamma/2$ with a width $\Gamma' = (\sqrt{3} - 1)(\Gamma/2)$.

Several other spectral features need to be resolved as well, aside from the sharply peaked resonances. The density of states around the conduction band edge

varies sharply with respect to energy. We have optimized this turn-on using three different distributions: 1) a $1/\sqrt{E}$ distribution convolved with a Lorentzian, 2) a triangular distribution, and 3) a homogeneous distribution around the conduction band edge. We find that the homogeneous contribution around the conduction band edge provides an adequate resolution of the density of states at the conduction band edge. The integral contributions to the density of states at the band edges is modeled by the monotonic function $m_{band-edge}$. The variations in the Fermi occupation factor, f, in the contacts are resolved in a similar manner noting that $m_{fermi} = 1 - f$ is an analytic and monotonically growing function. To resolve turn-ons of the Fermi distribution we optimize the energy grid with respect to the fourth derivative of f resulting in another monotonic contribution m_{fermi}. Given the individual contributions we can build an analytic monotonic function $m(E) = x_1 m_{IM}(E) + x_2 m_{RE}(E) + x_3 m_{RE-IM}(E) + x_4 m_{band-edge} + x_5 m_{fermi} + x_6 m_{fermi}$, where x_i are weighing parameters. The monotonic function $m(E)$ can be evaluated "cheaply" for any E and can be inverted to a homogeneous grid in $y = m(E)$ numerically.

As a comparison we perform here the integration over the real and the imaginary part of 2 narrow resonances ($\Gamma = 0.5\mu eV$) at $E_{r_1} \simeq 100\,meV$ and $E_{r_2} \simeq 300\,meV$ placed in a wide energy range of $1000\,meV$ (see inset of figure (2a)). We perform the numerical integration in four different schemes: 1) our new fixed energy grid with trapezoidal integration (see Fig. (2a)), 2) our new fixed energy grid with Simpson integration (see Fig. (2b,c)), 3) a typical adaptive grid scheme with trapezoidal integration (see Fig. (2a,b)), and 4) a typical adaptive grid scheme with Simpson integration (see Fig. (2c)). Figure (2) shows clearly that we typically obtain a savings of integration nodes of a factor of $k \approx 4 - 10$. Depending on the calculation that has to be performed the number of operations may scale as $k^1 - k^3$ and the required memory scales as $k^1 - k^2$.

The drawback of the adaptive schemes are that they have to "home in" on the resonances, which pushes the number of integration nodes up. The drawback of the fixed energy grid scheme described here is that we
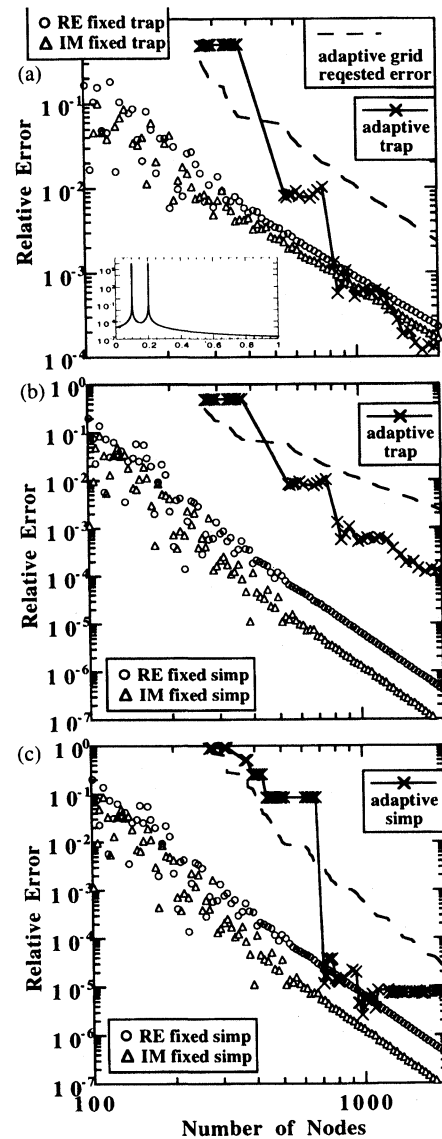

FIGURE 2 Demonstration of the efficiency of our new fixed energy grid compared to a "standard" adaptive grid approach. We integrate the real and imaginary part of two resonances located at about 100 meV and 300 meV, respectively, each with an energy width of 0.5μeV. The energy range of integration is 0-1eV (see inset in a). (a) trapezoidal integration using the fixed and the adaptive grid. The dashed line indicates the requested relative accuracy of the adaptive grid, which is clearly different from the absolute accuracy. (b) Simpson integration using the fixed grid compared to trapezoidal integration on an adaptive grid. (c) Simpson integration with the fixed grid and the adaptive grid. For the fixed grid we show the relative errors of the imaginary part and the real part integral separately. For the adaptive grid approaches we show the maximum error. The dashed lines indicate the relative precision specified in the adaptive grid approaches

have to perform the resonance finding first. Single band resonance finding [2] scales as N^2 and multi-band resonance finding [3] scales superlinear in N where N is the number of tight binding spatial nodes in the device. The recursive Green function algorithm [4] which can deliver the density of states and the transmission throughout the device scales as N. For the modeling of coherent transport [5] with not too sharp resonances ($\Gamma > 1\mu eV$) we find the adaptive grid optimization outperformes the fixed energy grid scheme including resonance finding. However in cases where the resonances are very sharp ($\Gamma < 1\mu eV$) our new scheme outperforms the usual adaptive scheme. For the modeling of incoherent transport [6] where we include acoustic and polar optical phonons, alloy disorder scattering and interface roughness scattering we find our fixed energy grid scheme essential. In this case we have to perform multidimensional, coupled integrals, where intermediate results have to be stored and we find ourselves severely limited by memory requirements of 400-500MB on a workstation. Another drawback of the fixed grid scheme presented here is that the integration accuracy cannot be specified a priori. Note, however, that the relative accuracy specified in the standard adaptive schemes does not correspond to the absolute accuracy (Fig. 2). The standard adaptive scheme can easily miss resonances completely and it often turns out to be an "all or nothing approach". With respect to accuracy we sometimes find it useful to mix the two approaches, where the fixed grid serves as a starting point of adaptive grid integration.

Acknowledgements

References

[1] W. H. Press, B. P. Flannery, A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1987), p. 114.
[2] C. L. Fernando and W. R. Frensley, J. Appl. Phys. **76**, 2881 (1994).
[3] R. C. Bowen, W. R. Frensley, G. Klimeck, and R. K. Lake, Phys. Rev. B **52**, 2754 (1995).
[4] J. A. Støveneng and E. H. Hauge, Phys. Rev. B **44**, 13582 (1991).
[5] G. Klimeck *et al.,* Appl. Phys. Lett. **67**, 2539 (1995).
[6] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, Journal of Applied Physics, **81**, 7845 (1997).

Biographies

Gerhard Klimeck performed this work as a post-doc at the University of Texas at Dallas and later as a Member Technical Staff at Texas Instruments Incorporated. He will move to the Jet Propulsion Laboratory in early 1998. His research interest is the modeling of electron transport in nanoelectronic devices.

Roger Lake is a Member Technical Staff at Raytheon TI Systems. He is working on scattering, disorder, and optical processes in III-V and Silicon quantum well semiconductor devices.

R. Chris Bowen performed this work as a graduate student at the University of Texas at Dallas and he is presently Member Technical Staff at Texas Instruments Incorporated. His current research interest is full-band modeling of quantum electron transport in Silicon and III-V structures.

Chenjing L. Fernando is a member technical staff at Texas Instruments Incorporated. Her present research interest is process synthesis.

William R. Frensley is the Head of the Electrical Engineering Program at the University of Texas at Dallas. His research interest is the development of interactive high performance heterostructure device simulators.